Lecture 4: Stationary Distributions

Last Time

Transition matrix : P

Initial data : $\vec{\mu} = (\mu_1, \cdots, \mu_{|S|})$, where

$\vec{\mu}_j = \mathbb{P}(X_0 = i_j)$, $\mathbb{P}(X_1 = i_k) = (\vec{\mu}P)_k$, and

$\mathbb{P}(X_n = i_k) = (\vec{\mu} \cdot P^n)_k$

Suppose $\mathbb{P}(X_0 = x) = \begin{cases} 1 , & \text{if } x = i_k ; \\ 0 & \text{else} . \end{cases}$

Then the initial data $\vec{\mu} = (0, \cdots, 0, 1, 0, \cdots, 0) = \delta_{i_k}$.

" k-th" entry

" Dirac mass @ i_k ".

Today

1°. Q: Suppose the chain starts at i_k , i.e. , $\vec{\mu} = \delta_{i_k}$,

What is the long time behaviour?

Let $\vec{\mu}(n)$ be the distribution of X_n , i.e.,

$[\vec{\mu}(n)]_j = \mathbb{P}(X_n = i_j)$. Then what is

$$\lim_{n\to\infty} \vec{\mu}(n) = ?$$

Prepared by Kelvin Shuangjian ZHANG (ks3zhang@waterloo.ca)

Ex1. (2 coffee shops)

Suppose there are two coffee shops: Tim Hortons and Starbucks. The following Markov chain describes the probability of going to each coffee shop given which coffee shop is visited last time.

$$P = \begin{array}{c c} & \begin{array}{cc} T & S \end{array} \\ \begin{array}{c} T \\ S \end{array} & \begin{bmatrix} 0.8 & 0.2 \\ 0.6 & 0.4 \end{bmatrix} \end{array}$$

Q: What is the long time behaviour?

A: $[\vec{\mu}(n)]_T = [\vec{\mu} P^n]_T = \vec{\mu}_T [P^n]_{TT} + \vec{\mu}_S [P^n]_{ST}$,

From the left figures, we know

$$\lim_{n\to\infty} [P^n]_{TT} = \lim_{n\to\infty} [P^n]_{ST} = 0.75,$$

and $$\lim_{n\to\infty} [P^n]_{SS} = \lim_{n\to\infty} [P^n]_{TS} = 0.25.$$

Thus, $$\lim_{n\to\infty} [\vec{\mu}(n)]_T = 0.75(\vec{\mu}_T + \vec{\mu}_S) = 0.75,$$

and $$\lim_{n\to\infty} [\vec{\mu}(n)]_S = \lim_{n\to\infty} (1 - [\vec{\mu}(n)]_T) = 0.25.$$

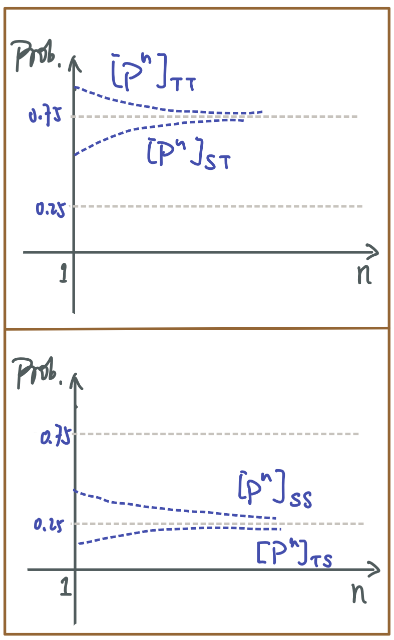

Fig 1

Remark 1. Suppose the limit exists and $\lim_{n\to\infty} \vec{\mu}(n) = \vec{\pi}$.

Then taking limits at both sides of

$$\vec{\mu}(n) = \vec{\mu}(n-1) \cdot P$$

yields $\vec{\pi} = \vec{\pi} P$.

That is, the limit is a left eigenvector of P with eigenvalue 1.

Def (Stationary distribution)

A probability vector $\vec{\pi}$ (i.e., $\pi_i \geq 0, \forall i$; $\sum_i \pi_i = 1$) is a stationary distribution of P if

$$\boxed{\vec{\pi} P = \vec{\pi}.}$$

Def. (Stationary measure)

A measure $\vec{\nu}$ on $\mathcal{X}$ (i.e., $\vec{\nu} \in \mathbb{R}^{|\mathcal{X}|}$, s.t. $\vec{\nu}_i \geq 0, \forall i$; $\sum_i \vec{\nu}_i > 0$) is a stationary measure of P if

$$\boxed{\vec{\nu} P = \vec{\nu}.}$$

Prepared by Kelvin Shuangjian ZHANG (ks3zhang@waterloo.ca)

Ex1 (cont.) Suppose the limit exists and $\lim_{n\to\infty} \vec{\mu}(n) = \vec{\pi}$.

Solving the system of equations with positivity constraints

$$\begin{cases} \vec{\pi} P = \vec{\pi} ; \\ \pi_i \geq 0 , \quad i=1,2 ; \\ \pi_1 + \pi_2 = 1 . \end{cases}$$

yields $\vec{\pi} = (0.75, 0.25)$ as the unique solution. Thus $\vec{\mu}(n) \xrightarrow{n\to\infty} (0.75, 0.25)$.

Here, $(0.75, 0.25)$ is the unique stationary distribution of P. For any $c>0$, $c\vec{\pi} = (0.75c, 0.25c)$ is a stationary measure of P.

Ex2. SRW on $\mathbb{Z}_4$.

$$P = \begin{array}{c} \begin{array}{cccc} [0] & [1] & [2] & [3] \end{array} \\ \begin{bmatrix} 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \end{bmatrix} \end{array} \begin{array}{c} \\ [0] \\ [1] \\ [2] \\ [3] \end{array}$$

bi-stochastic

Let $\vec{\nu} = (1, 1, 1, 1)$. Then $\vec{\nu} P = \vec{\nu}$.

Let $\vec{\pi} = c\vec{\nu}$ with $c = \frac{1}{4}$, then $\sum_i \vec{\pi}_i = c \sum_i \vec{\nu}_i = 1$.

Thus, $\vec{\pi} = (\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$ is a stationary distribution of P.

Remark 2. In general, if $|\mathcal{X}| < \infty$, and $\vec{\nu}$ is a stationary measure, then $\dfrac{\vec{\nu}}{\sum_i \vec{\nu}_i}$ is a stationary distribution.

Ex3. SRW on $\mathbb{Z}$.

$$P = \begin{array}{cc} & \begin{array}{ccccccc} \cdots & -2 & -1 & 0 & 1 & 2 & \cdots \end{array} \\ \begin{array}{c} \vdots \\ -2 \\ -1 \\ 0 \\ 1 \\ 2 \\ \vdots \end{array} & \begin{bmatrix} & \vdots & \vdots & \vdots & \vdots & \vdots & \\ \cdots & 0 & \frac{1}{2} & 0 & 0 & 0 & \cdots \\ \cdots & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & \cdots \\ \cdots & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & \cdots \\ \cdots & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & \cdots \\ \cdots & 0 & 0 & 0 & \frac{1}{2} & 0 & \cdots \\ & \vdots & \vdots & \vdots & \vdots & \vdots & \end{bmatrix} \end{array}$$

Let $\vec{\nu} = (\cdots, 1, 1, 1, 1, 1, \cdots)$, then

$$[\vec{\nu} P]_j = \sum_i \vec{\nu}_i P_{ij} = \sum_i P_{ij} = \frac{1}{2} + \frac{1}{2} = \vec{\nu}_j, \quad \forall j \in \mathbb{Z}.$$

This implies $\vec{\nu} P = \vec{\nu}$.

Notice that $\sum_i \vec{\nu}_i = \infty$ and thus it cannot be normalize into a probability vector.

Q: Does P have a stationary distribution?

2°. Proposition of Stationary distributions.

For $X_0 \sim \vec{\pi}$, then $X_t \sim \vec{\pi}$, for all $t \in \mathbb{N}$.

Key Questions:

Q1. When do stationary distributions exist?

Q2. Under Q1, when are they unique?

Q3. Under Q1 and Q2, suppose there exists a unique stationary distribution $\vec{\pi}$, let $\vec{\mu}(n)$ be the distribution of X_n under initial data $X_0 \sim \vec{\mu}$, will it always true that

$$\lim_{n \to \infty} \vec{\mu}(n) = \vec{\pi} \quad ?$$

Prepared by Kelvin Shuangjian ZHANG (ks3zhang@waterloo.ca)

3°. Detailed Balance condition:

Def. (Detailed Balance Condition)

$\vec{\pi}$ is said to satisfy the detailed balance condition if

$$\boxed{\vec{\pi}_x P_{xy} = \vec{\pi}_y P_{yx}, \quad \forall x, y \in \mathcal{X}.}$$

Lemma 1: If $\vec{\pi}$ satisfies the detailed balance condition, then $\vec{\pi}$ is sationary.

Pf. For any $y \in \mathcal{X}$,

$$\begin{aligned}[\vec{\pi}P]_y &= \sum_{x\in\mathcal{X}} \vec{\pi}_x P_{xy} \\ &= \sum_{x\in\mathcal{X}} \vec{\pi}_y P_{yx} \\ &= \vec{\pi}_y \sum_{x\in\mathcal{X}} P_{yx} \\ &= \vec{\pi}_y \cdot 1 \\ &= \vec{\pi}_y.\end{aligned}$$

Thus, $\vec{\pi}P = \vec{\pi}$.

Prepared by Kelvin Shuangjian ZHANG (ks3zhang@waterloo.ca)

Remark 3. (Interpretation of the detailed balance condition).

Given the initial data $X_0 \sim \vec{\pi}$, then $\forall x, y \in \mathcal{X}$,

$$\begin{aligned}
&\mathbb{P}(X_0 = x, X_1 = y) \\
&= \mathbb{P}(X_1 = y \mid X_0 = x) \cdot \mathbb{P}(X_0 = x) \\
&= P_{xy} \cdot \vec{\pi}_x \\
&= P_{yx} \cdot \vec{\pi}_y \\
&= \mathbb{P}(X_1 = x \mid X_0 = y) \cdot \mathbb{P}(X_0 = y) \\
&= \mathbb{P}(X_0 = y, X_1 = x).
\end{aligned}$$

This is the end of this lecture!

Prepared by Kelvin Shuangjian ZHANG (ks3zhang@waterloo.ca)